

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:48:50 ; Search time 16 Seconds
(without alignments)
2392.666 Million cell updates/sec

Title: US-09-497-822c-19
Perfect score: 4912
Sequence: 1 MEVQLGLGRVPRPPSKTYR.....SVQVPKILSKVKPIYFHTQ 923

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4880	99.3	919	1	ANDR_HUMAN
2	4794	97.6	911	1	ANDR_PANTR
3	4694	95.6	895	1	ANDR_PAPHA
4	4678	95.2	895	1	ANDR_MACFA
5	4404.5	89.7	884	1	ANDR_EULFC
6	4319	87.9	907	1	ANDR_CANFA
7	4200.5	85.5	902	1	ANDR_RAT
8	4176	85.0	899	1	ANDR_MOUSE
9	3509.5	71.4	709	1	ANDR_RABIT
10	1262.5	25.7	930	1	PRGR_RABIT
11	1256	25.6	933	1	PRGR_HUMAN
12	1234	25.1	923	1	PRGR_RAT
13	1218	24.8	786	1	PRGR_CHICK
14	1205	24.5	923	1	PRGR_MOUSE
15	1116	22.7	377	1	PRGR_SHEEP
16	1092	22.2	981	1	MCR_RAT
17	1078.5	22.0	984	1	MCR_HUMAN
18	1069	21.8	795	1	GCR_RAT
19	1063.5	21.7	783	1	GCR_MOUSE
20	1053.5	21.4	776	1	GCR_XENLA
21	1053.5	21.4	777	1	GCR_SAGOE
22	1046.5	21.3	777	1	GCR_AOTNA
23	1042	21.2	777	1	GCR_HUMAN
24	1042	21.2	778	1	GCR_SAISC
25	1041.5	21.2	777	1	GCR_SAIBB
26	1034	21.1	612	1	MCR_XENLA
27	1033.5	21.0	776	1	GCR_TUPGB
28	1023	20.8	758	1	GCR_ONCMY
29	1014	20.6	977	1	MCR_TUPGB
30	1005	20.5	807	1	GCR_FAROL
31	996	20.3	771	1	GCR_CAVPO
32	970.5	19.8	703	1	GCR_PIG
33	509	10.4	180	1	PRGR_MACEU

34	502	10.2	595	1	ESR1_PIG	Q29040 sus scrofa
35	498.5	10.1	594	1	ESR1_HORSE	Q91v98 equus cabal
36	491	10.0	586	1	ESR1_XENLA	P81559 xenopus lae
37	485.5	9.9	600	1	ESR1_RAT	P06211 rattus norv
38	481	9.8	595	1	ESR1_HUMAN	P03372 homo sapien
39	474.5	9.7	595	1	ESR1_MESAU	Q9qz15 mesocricetu
40	474	9.6	589	1	ESR1_CHICK	F06212 gallus gall
41	474	9.6	599	1	ESR1_MOUSE	P19785 mus musculu
42	470.5	9.6	569	1	ESR1_BRARE	P57717 brachydanio
43	469.5	9.6	622	1	ESR1_ONCMY	P16058 oncorhynchu
44	465	9.5	587	1	ESR1_POEGU	Q1250 poephila gu
45	461.5	9.4	620	1	ESR1_ORVLA	P30241 cryzias lat

ALIGNMENTS

RESULT 1

ID	ANDR_HUMAN	STANDARD;	PRT;	919 AA.
AC	P10275;			
DT	01-MAR-1989 (Rel. 10, Created)			
DT	01-APR-1990 (Rel. 14, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Androgen receptor (Dihydrotestosterone receptor).			
GN	AR OR NR3C4 OR DHTR.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
EX	MEDLINE=89112208; PubMed=3216866;			
RA	Lubahn D.B., Joseph D.R., Sar M., Tan J., Higgs H.N., Larson R.E.,			
RA	French F.S., Wilson E.M.;			
RT	"The human androgen receptor: complementary deoxyribonucleic acid			
RT	cloning, sequence analysis and gene expression in prostate."			
RL	Mol. Endocrinol. 2:1265-1275(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A. AND VARIANT CALS MET-866.			
EX	MEDLINE=90083302; PubMed=2594783;			
RA	Lubahn D.B., Brown T.R., Simental J.A., Higgs H.N., Migeon C.J.,			
RA	Wilson E.M., French F.S.;			
RT	"Sequence of the intron/exon junctions of the coding region of the			
RT	human androgen receptor gene and identification of a point mutation			
RT	in a family with complete androgen insensitivity."			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:9534-9538(1989).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
EX	MEDLINE=90258935; PubMed=2342476;			
RA	Govindan M.V.;			
RT	"Specific region in hormone binding domain is essential for hormone			
RT	binding and trans-activation by human androgen receptor."			
RL	Mol. Endocrinol. 4:417-427(1990).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
EX	MEDLINE=90258935; PubMed=3174628;			
RA	Chang C., Kokontis J., Liao S.;			
RT	"Structural analysis of complementary DNA and amino acid sequences of			
RT	human and rat androgen receptors."			
RL	Proc. Natl. Acad. Sci. U.S.A. 85:7211-7215(1988).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
EX	TISSUE=Prostate;			
RA	MEDLINE=89017168; PubMed=2911578;			
RT	Tilley W.D., Marcelli M., Wilson J.D., McPhaul M.J.;			
RT	"Characterization and expression of a cDNA encoding the human androgen			
RT	receptor."			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:327-331(1989).			
RN	[6]			
RP	SEQUENCE FROM N.A.			
EX	TISSUE=Prostate;			

Fri May 9 14:20:44 2003

- RX MEDLINE=91155943; PubMed=2293020;
RA Marcelli M., Tilley W.D., Wilson C.M., Griffin J.E., Wilson J.D.,
RA McPhaul M.J.;
RT "Definition of the human androgen receptor gene structure permits the
RT identification of mutations that cause androgen resistance: premature
RT termination of the receptor protein at amino acid residue 588 causes
RT complete androgen resistance.";
RL Mol. Endocrinol. 4:1105-1116(1990).
RN [7]
RN SEQUENCE OF 189-919 FROM N.A.
RP MEDLINE=8817811; PubMed=3353726;
RX Chang C., Kokontis J., Liao S.;
RA "Molecular cloning of human and rat complementary DNA encoding
RT androgen receptors.";
RT Science 240:324-326(1988).
RN [8]
RN SEQUENCE OF 468-919 FROM N.A.
RP MEDLINE=88240407; PubMed=3377788;
RX Trapman J., Klaassen P., Kuiper G.G.J.M., van der Korp J.A.G.M.,
RA Faber P.W., van Rooij H.C.J., Geurts van Kessel A., Voorhorst M.M.,
RA Mulder E., Brinkmann A.O.;
RT "Cloning, structure and expression of a cDNA encoding the human
RT androgen receptor.";
RL Biochem. Biophys. Res. Commun. 153:241-248(1988).
RN [9]
RN INTERACTION WITH RAN.
RX MEDLINE=99329028; PubMed=10400640;
RA Hsiao P.-W., Lin D.-L., Nakao R., Chang C.;
RT "The linkage of Kennedy's neuron disease to ARA24, the first
RT identified androgen receptor polyglutamine region-associated
RT coactivator.";
RL J. Biol. Chem. 274:20229-20234(1999).
RN [10]
RN POLYMORPHISM OF POLY-GLN REGION.
RX MEDLINE=92220629; PubMed=1561105;
RA Sledzgens H.P.B.M., Oostra B.A., Brinkmann A.O., Trapman J.;
RT "Trinucleotide repeat polymorphism in the androgen receptor gene
RT (AR).";
RL Nucleic Acids Res. 20:1427-1427(1992).
RN [11]
RN POLYMORPHISM OF POLY-GLY REGION.
RC TISSUE-Blood;
RL Lu J., Daniels M.;
RA Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
RN [12]
RN POLYMORPHISM OF POLY-GLN REGION.
RX MEDLINE=97250535; PubMed=9096391;
RA Giovannucci E., Stampfer M.J., Krithivas K., Brown M., Brufsky A.,
RA Talcott J., Hennekens C.H., Kantoff P.W.;
RT "The CAG repeat within the androgen receptor gene and its
RT relationship to prostate cancer.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:3320-3323(1997).
RN [13]
RN ERRATUM.
RA Giovannucci E., Stampfer M.J., Krithivas K., Brown M., Dahl D.,
RA Brufsky A., Talcott J., Hennekens C.H., Kantoff P.W.;
RL Proc. Natl. Acad. Sci. U.S.A. 94:8272-8272(1997).
RN [14]
RN REVIEW ON VARIANTS.
RX MEDLINE=93092459; PubMed=1458719;
RA Pinsky L., Trifiro M.A., Kaufman M., Beitel L.K., Mhatre A.,
RA Kazemi-Esfarjani P., Sabbaghian N., Lumbroso R., Alvarado C.,
RA Vasiliou M., Gottlieb B.;
RT "Androgen resistance due to mutation of the androgen receptor.";
RL Clin. Invest. Med. 15:456-472(1992).
RN [15]
RN REVIEW ON VARIANTS AIS.
RX MEDLINE=93339360; PubMed=8339746;
RA Brown T.R., Scherer P.A., Chang Y.-T., Migeon C.J., Ghirri P.,
RA Muroto K., Zhou Z.;
RT "Molecular genetics of human androgen insensitivity.";
RL Eur. J. Pediatr. 152 suppl. 2:S62-S69(1993).
RN [16]
RN REVIEW ON VARIANTS.
RX MEDLINE=94059770; PubMed=8240973;
RA Sultan C., Lumbroso S., Poujol N., Belon C., Boudon C.,
RA Lobaccaro J.-M.;
RT "Mutations of androgen receptor gene in androgen insensitivity
RT syndromes.";
RL J. Steroid Biochem. Mol. Biol. 46:519-530(1993).
RN [17]
RN REVIEW ON VARIANTS.
RP MEDLINE=95023089; PubMed=7937057;
RX Patterson M.N., Hughes I.A., Gottlieb B., Pinsky L.;
RA "The androgen receptor gene mutations database.";
RL Nucleic Acids Res. 22:3560-3562(1994).
RN [18]
RN REVIEW ON VARIANTS.
RP MEDLINE=95352489; PubMed=7626493;
RX Brinkmann A.O., Jenster G., Ris-Stalpers C., van der Korp J.A.G.M.,
RA Bruggenwirth H.T., Boehmer A.L.M., Trapman J.;
RT "Androgen receptor mutations.";
RL J. Steroid Biochem. Mol. Biol. 53:443-448(1995).
RN [19]
RN REVIEW ON VARIANTS.
RP MEDLINE=97169385; PubMed=9016528;
RX Gottlieb B., Trifiro M.A., Lumbroso R., Vasiliou M., Pinsky L.;
RA "The androgen receptor gene mutations database.";
RL Nucleic Acids Res. 25:158-162(1997).
RN [20]
RN VARIANT LNCAP ALA-877.
RX MEDLINE=91083633; PubMed=2260966;
RA Veldscholte J., Ris-Stalpers C., Kuiper G.G.J.M., Jenster G.,
RA Berrevoets C.A., Claassen E., van Rooij H.C.J., Trapman J.,
RA Brinkmann A.O., Mulder E.;
RT "A mutation in the ligand binding domain of the androgen receptor of
RT human LNCap cells affects steroid binding characteristics and
RT response to anti-androgens.";
RL Biochem. Biophys. Res. Commun. 173:534-540(1990).
RN [21]
RN VARIANTS CAIS CYS-774: GLN-831 AND MET-866.
RX MEDLINE=91186983; PubMed=2082179;
RA Brown T.R., Lubahn D.B., Wilson E.M., French F.S., Migeon C.J.,
RA Corfen J.L.;
RT "Functional characterization of naturally occurring mutant androgen
RT receptors from subjects with complete androgen insensitivity.";
RL Mol. Endocrinol. 4:1759-1772(1990).
RN [22]
RN VARIANT CYS-774.
RX MEDLINE=91310758; PubMed=1856263;
RA Marcelli M., Tilley W.D., Zoppi S., Griffin J.E., Wilson J.D.,
RA McPhaul M.J.;
RT "Androgen resistance associated with a mutation of the androgen
RT receptor at amino acid 772 (Arg->Cys) results from a combination of
RT decreased messenger ribonucleic acid levels and impairment of
RT receptor function.";
RL J. Clin. Endocrinol. Metab. 73:318-325(1991).
RN [23]
RN VARIANT CAIS PRO-617.
RX MEDLINE=91154385; PubMed=1999491;
RA Marcelli M., Zoppi S., Grino P.B., Griffin J.E., Wilson J.D.,
RA McPhaul M.J.;
RT "A mutation in the DNA-binding domain of the androgen receptor gene
RT causes complete testicular feminization in a patient with
RT receptor-positive androgen resistance.";
RL J. Clin. Invest. 87:1123-1126(1991).
RN [24]
RN VARIANT PAIS CYS-763.
RX MEDLINE=91185626; PubMed=2010552;
RA McPhaul M.J., Marcelli M., Tilley W.D., Griffin J.E.,
RA Isidro-Gutierrez R.F., Wilson J.D.;
RT "Molecular basis of androgen resistance in a family with a qualitative
RT abnormality of the androgen receptor and responsive to high-dose
RT androgen therapy.";
RL J. Clin. Invest. 87:1413-1421(1991).
RN [25]

[illegible]


```
QY 241 LKAVSVSMGLGVLEALHSPGQLRGDCMYAPLLGVPPAVRPTPCAPLAECKGSLDD 300
Db 222 LKAVSVSMGLGVLEALHSPGQLRGDCMYAPLLGVPPAVRPTPCAPLAECKGSLDD 281
QY 301 AGKSTEDTAESYSPKGGYTKLGESLGCSSAAGSSGTLELPSTLSLYKSGALDAAA 360
Db 282 AGKSTEDTAESYSPKGGYTKLGESLGCSSAAGSSGTLELPSTLSLYKSGALDAAA 341
QY 361 YQSDRYNFPALAGPPPPPPHAPHARIKENLDYGSAAWAAAAACRYGDLASLHGAG 420
Db 342 YQSDRYNFPALAGPPPPPPHAPHARIKENLDYGSAAWAAAAACRYGDLASLHGAG 401
QY 421 AAGPGSGSPSAASSSSWHTLFTABEGQLYPCGGGGGGGGGGGGGGGGGGGGG 480
Db 402 AAGPGSGSPSAASSSSWHTLFTABEGQLYPCGGGGGGGGGGGGGGGGGGGGG 462
QY 481 VAPGYTRPPGGLAGQSDFTAPDWWYPGMVSRVPPSPCTCYKSEMPWDSYSGPYGD 540
Db 453 VAPGYTRPPGGLAGQSDFTAPDWWYPGMVSRVPPSPCTCYKSEMPWDSYSGPYGD 512
QY 541 MRLETARDHVLPIIDYPPQKTKLICGDEASGCHYGALTCGCKVFFKRAEKGKYLCA 600
Db 513 MRLETARDHVLPIIDYPPQKTKLICGDEASGCHYGALTCGCKVFFKRAEKGKYLCA 572
QY 601 SRNCTIDKFRKNCPSRLRKYEAGMTLGARKLKLGNLKLQEGEASSTTSPTBEET 660
Db 573 SRNCTIDKFRKNCPSRLRKYEAGMTLGARKLKLGNLKLQEGEASSTTSPTBEET 632
QY 661 QKLVSHIEGECOPILNVLEATEPGVVCAGHDNDNPDFAALLSLNLGERQLVHV 720
Db 633 QKLVSHIEGECOPILNVLEATEPGVVCAGHDNDNPDFAALLSLNLGERQLVHV 692
QY 721 KWAKALPGFRLNHDVDDQMAVIOYSWGMGLVFMAGWRSFTNVNSRMILYFADLVNEYRMH 780
Db 693 KWAKALPGFRLNHDVDDQMAVIOYSWGMGLVFMAGWRSFTNVNSRMILYFADLVNEYRMH 752
QY 781 KSRMYSQCVNRHLSEFGWLIQTPQEFCLMKALLLSIIPVGLKNQKFFDELRMNLIK 840
Db 753 KSRMYSQCVNRHLSEFGWLIQTPQEFCLMKALLLSIIPVGLKNQKFFDELRMNLIK 812
QY 841 ELDRILACKRNKPTSCSRFFYQLTKLDSVQPTARELHQFTDLILKSHMVSDFFEMMA 900
Db 813 ELDRILACKRNKPTSCSRFFYQLTKLDSVQPTARELHQFTDLILKSHMVSDFFEMMA 872
QY 901 ELISVQVPKILSGKVKPIYFHTQ 923
Db 873 ELISVQVPKILSGKVKPIYFHTQ 895
```

```
RESULT 4
ANDR_MACFA STANDARD: PRT; 895 AA.
AC 097952;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Androgen receptor (Dihydrotestosterone receptor).
GN AR OR NR3C4.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
FN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98404153; PubMed=9732460;
RA Choong C.S., Kempainen J.A., Wilson E.M.;
RT "Evolution of the primate androgen receptor: a structural basis for
RL disease.";
RL J. Mol. Evol. 47:334-342(1998).
CC -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
```

```
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC NR3 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U94179; AAC73050.1; -
CC HSP: P06536; IGDC.
CC TRANSFAC: T04654; -
CC InterPro: IPR001103; Andrgn_receptor.
CC InterPro: IPR000536; Hormone_rec_lig.
CC InterPro: IPR001628; Znf_C4steroid.
CC Pfam: PF00104; hormone_rec; 1.
CC Pfam: PF00105; zf-C4; 1.
CC Pfam: PF02166; Androgen_recep; 1.
CC PRINTS: PR00047; STROIDFINGER.
CC ProDom: PD000035; Znf_C4steroid; 1.
CC SMART: SM00430; HOLI; 1.
CC SMART: SM00399; Znf_C4; 1.
CC PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
CC Receptor; Transcription regulation; DNA-binding; Nuclear protein;
CC Zinc-finger; Steroid-binding.
CC MODULATING (BY SIMILARITY):
CC DOMAIN 1 533 NUCLEAR RECEPTOR-TYPE.
CC DNA_BIND 535 600 C4-TYPE.
CC ZN_FING 535 555 C4-TYPE.
CC ZN_FING 571 595 C4-TYPE.
CC DOMAIN 666 895 LIGAND-BINDING.
CC DOMAIN 55 62 POLY-GLN.
CC DOMAIN 68 74 POLY-GLN.
CC DOMAIN 178 182 POLY-GLN.
CC DOMAIN 357 366 POLY-PRO.
CC DOMAIN 381 387 POLY-ALA.
CC DOMAIN 434 448 POLY-GLY.
CC SEQUENCE 895 AA; 96494 MW; A3EB17916F43A097 CRC64;
CC -----
Query Match 95.2%; Score 4678; DB 1; Length 895;
Best Local Similarity 95.9%; Pred. No. 6.8e-216;
Matches 886; Conservative 2; Mismatches 6; Indels 30; Gaps 3;
QY 1 MEVQLGLGRVYPRPSKTYRGAFQNLFGSVREVIONPGRHPPEAASAPPGASLILLQQQ 60
Db 1 MEVQLGLGRVYPRPSKTYRGAFQNLFGSVREVIONPGRHPPEAASAPPGASL 54
QY 61 QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ 119
Db 55 -----QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ 100
QY 120 QPQSALECHPERGCVPEPGAAVAASGLPQQLPAPPPEDDDSAAPSTLSLLGFTFFGLSSC 179
Db 101 QPQSALECHPERGCVPEPGAAVAAGKGLPQQLPAPPPEDDDSAAPSTLSLLGFTFFGLSSC 160
QY 180 SADLKDILSEASTMQLQQQQQEAIVSESSSSGRAREASGAPTSKDNVILGTTSTISDNK 239
Db 161 STDLDKILSEASTMQLQQQQQEAIVSESSSSGRAREASGAPTSKDNVILGTTSTISDNK 220
QY 240 ELCKAVSVSMGLGVLEALHSPGQLRGDCMYAPLLGVPPAVRPTPCAPLAECKGSLDD 299
Db 221 ELCKAVSVSMGLGVLEALHSPGQLRGDCMYAPLLGVPPAVRPTPCAPLAECKGSLDD 280
QY 300 SAGKSTEDTAESYSPKGGYTKLGESLGCSSAAGSSGTLELPSTLSLYKSGALDEAA 359
Db 281 SAGKSTEDTAESYSPKGGYTKLGESLGCSSAAGSSGTLELPSTLSLYKSGALDEAA 340
QY 360 AYQSDRYNFPALAGPPPPPPHAPHARIKENLDYGSAAWAAAAACRYGDLASLHGA 419
Db 342 AYQSDRYNFPALAGPPPPPPHAPHARIKENLDYGSAAWAAAAACRYGDLASLHGA 401
```



```
QY 540 DMRLETARDHVLPIIDYPPQKTCICGDEASGCHYGALTGCGSKVFFKRAAEKQKYL 599
Db 501 DMRLETARDHVLPIIDYPPQKTCICGDEASGCHYGALTGCGSKVFFKRAAEKQKYL 560
QY 600 ASRNDCTIDKFRKNCPCRLKCKVEAGMTLGARKLKLGNLKLQEGEASSTSPTEET 659
Db 561 ASRNDCTIDKFRKNCPCRLKCKVEAGMTLGARKLKLGNLKLQEGEASSTSPTEES 620
QY 660 TQKLTSHIEGYECQIFLNVLEATEPGVWCAGHNNQPDSSFAALLSSNLGELGERQLVHV 719
Db 621 SOKLTSHIEGYECQIFLNVLEATEPGVWCAGHNNQPDSSFAALLSSNLGELGERQLVHV 680
QY 720 VKAKALPGERNLHVDDQMAVIOQSWGLMVFAMGWSFTNVNSRMLYFAPDLVFNERYM 779
Db 681 VKAKALPGERNLHVDDQMAVIOQSWGLMVFAMGWSFTNVNSRMLYFAPDLVFNERYM 740
QY 780 HKRMYSCVRRMHLQSGFWLQITPQEFCKMALLFSIIPVDGLKNQKFFDELNMNYI 839
Db 741 HKRMYSCVRRMHLQSGFWLQITPQEFCKMALLFSIIPVDGLKNQKFFDELNMNYI 800
QY 840 KELDRITACKKNTSGRRYQTLKLDVSQPIARELHQFTFDLLIKSHMVSVDPEMM 899
Db 801 KELDRITACKKNTSGRRYQTLKLDVSQPIARELHQFTFDLLIKSHMVSVDPEMM 860
QY 900 AEILSVQPKILSGKVPIVPHQ 923
Db 861 AEILSVQPKILSGKVPIVPHQ 884

RESULT 6
ID ANDR_CANFA STANDARD; PRT; 907 AA.
AC Q9T90;
DT 30-MAY-2000 (Rel. 39, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Androgen receptor (Dihydrotestosterone receptor).
GN AR OR NR3C4.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21618348; PubMed-11768233;
RA Lu B., Smock S.L., Castleberry T.A., Owen T.A.;
RT "Molecular cloning and functional characterization of the canine
androgen receptor."
RL Mol. Cell. Biochem. 226:129-140(2001).
CC -1- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC NR3 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF197950; AAF18084.1;
CC HSSP; P06536; 1GDC.
CC InterPro; IPR001103; Andrgn_receptor.
CC InterPro; IPR000536; Hormone_rec_lig.
CC InterPro; IPR001628; Znf_C4steroid.
CC Pfam; PF00104; hormone_rec; 1.
```

```
DR Pfam; PF00105; zf-C4; 1.
DR Pfam; PF02166; Androgen_recep; 1.
DR PRINTS; PR00047; STROIDFINGER.
DR ProDom; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOLI; 1.
DR SMART; SM00399; Znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger; Steroid-binding.
KW DOMAIN 1 545 MODULATING (BY SIMILARITY).
FT DNA_BIND 547 612 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 547 567 C4-TYPE.
FT ZN_FING 583 607 C4-TYPE.
FT DOMAIN 678 907 LIGAND-BINDING.
FT DOMAIN 55 64 POLY-GLN.
FT DOMAIN 70 76 POLY-GLN.
FT DOMAIN 131 134 POLY-GLN.
FT DOMAIN 180 202 POLY-GLN.
FT DOMAIN 329 332 POLY-SER.
FT DOMAIN 375 384 POLY-PRO.
FT DOMAIN 399 405 POLY-ALA.
SQ SEQUENCE 907 AA; 98726 MW; C8619F78DD2338AF CRC64;

Query Match 87.9%; Score 4319; DB 1; Length 907;
Best Local Similarity 87.3%; Pred. No. 8.3e-199;
Matches 822; Conservative 20; Mismatches 46; Indels 54; Gaps 5;

QY 1 MEVQLGLGRVYPRPSKTYRGAFQNLFGSVREVIONPGRHPEAASAPPGASILLQLQQ 60
Db 1 MEVQLGLGRVYPRPSKTYRGAFQNLFGSVREVIONPGRHPEAASAPPGASILLQLQQ 54
QY 61 QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ 119
Db 55 -----QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ 102
QY 120 QPQSALECHPERGCVPEPCAANAASKGLPQQLPAPDEDDSAAPTSLILGPTFFGLSSC 179
Db 103 QQRSAKSGHPESACVPEPGVTSATKGLQQQQAPPDENDSAAPTSLILGPTFFGLSSC 162
QY 180 SADLKDILSEASTMOLL-----QQQQQEAQVSSSSSGRAREASGAPT 221
Db 163 STDLDKILSEAGTMOLLQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ 222
QY 222 SKDNLYLGTSTISDNAKELCKAVSMGLGVEALEHLSPEQLRGDCMYAPLLGVPPAV 281
Db 223 SSKDSYLGGSSTISDSAKELCKAVSMGLGVEALEHLSPEQLRGDCMYAPLLGVPPAV 282
QY 282 RPTPCAPLAECKGSLDSDSAGKSTEDTAESPFKGYGTYKLEGESLGCSSGSAAGSSTL 341
Db 283 R--PCAPLAECKGSLDSDGPKGTEETAESPFKAGYAKGLDGLDGLSGSSSEAGSGTL 340
QY 342 ELPSTLSLYKSGALDEAAAYQSRDYNFPLAGPPPPPPHPPHARLTKLENPLDYGSAW 401
Db 341 EMPSTLSLYKSGALDEAAAYQSRDYNFPLSGPPPPPPHPPHTRIKLENPLDYGSAW 400
QY 402 AAAAAQCRYGDLASLHGAGAGPGSGSPSAASSSWHTLFTAEQQLYCPGCGGGGGGG 461
Db 401 AAAAAQCRYGDLASLHGAGAGPGSGSPSATSSSWHTLFTAEQQLYCPGCGGGGGSGAG 460
QY 462 GGGGGGGGGGGGGGGAAGAPYGYTRPPQGLAGQESFTADPVWYPGGMVSRVYPSPT 521
Db 461 DG-----GSAVPGYGYTRPPQGLAGQEGDFPPDPVWYPGGVVSRVYPPSPS 505
QY 522 CVKSEMGPMWDSYSGPYGDMRLTARDHVLPIIDYFFPQKTCICGDEASCHVGLTGC 581
Db 506 CVKSEMGSMWESYSGPYGDMRLTARDHVLPIIDYFFPQKTCICGDEASCHVGLTGC 565
QY 582 SKVFFKRAAEKQKYLKASRNDCTIDKFRKNCPCRLKCKVEAGMTLGARKLKLGNL 641
Db 566 SKVFFKRAAEKQKYLKASRNDCTIDKFRKNCPCRLKCKVEAGMTLGARKLKLGNL 625
QY 642 KLOEGEASSTSPTEETQKLTSHIEGYECQIFLNVLEATEPGVWCAGHNNQPDSP 701
Db 642 KLOEGEASSTSPTEETQKLTSHIEGYECQIFLNVLEATEPGVWCAGHNNQPDSP 701
```


Db 626 KLOEGBEASNTSPTEPTQKLTVSHIEGYECQPIFNVLNLEAIEPGVVCAGHNNQPDSP 685
Qy 702 AALLSSNLGELQRLVHVYKAKALPGFRNLHVDQMAVIOYSWMGLVFMWGRSFTNV 761
Db 686 AALLSSNLGELQRLVHVYKAKALPGFRNLHVDQMAVIOYSWMGLVFMWGRSFTNV 745
Qy 762 NSRMLYFAPDLVFNRYMHKSRMTSQCVRMHLSQEFQWGLQITPQEFCKMALLFSIIP 821
Db 746 NSRMLYFAPDLVFNRYMHKSRMTSQCVRMHLSQEFQWGLQITPQEFCKMALLFSIIP 805
Qy 822 VDGLKNQKFFDELNMNLIKELDRILIAKCKRNPTSCSRFFQTLKLLDSVOPIARELHQFT 881
Db 806 VDGLKNQKFFDELNMNLIKELDRILIAKCKRNPTSCSRFFQTLKLLDSVOPIARELHQFT 865
Qy 882 FDLILKSHMVSVDPEMMAEIIISQVVKILSGKVKPIYFHTQ 923
Db 866 FDLILKSHMVSVDPEMMAEIIISQVVKILSGKVKPIYFHTQ 907
RESULT 7
ID ANDR_RAT STANDARD; PRT; 902 AA.
AC P15207; Q63049;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Androgen receptor (Dihydrotestosterone receptor).
GN AR OR NR3C4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID-10116;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE-89112209; PubMed-3216867;
RA Tan J., Joseph D.R., Quarby V.E., Lubahn D.B., Sar M., French F.S.,
RA Wilson E.M.;
RT "The rat androgen receptor: primary structure, autoregulation of its
RT messenger ribonucleic acid, and immunocytochemical localization of
RT the receptor protein.";
RL Mol. Endocrinol. 2:1276-1285(1988).
RN [2]
RN SEQUENCE FROM N.A.
RP TISSUE-Prostate;
RX MEDLINE-89017168; PubMed-3174628;
RA Chang C., Kokontis J., Ligo S.;
RT "Structural analysis of complementary DNA and amino acid sequences of
RT human and rat androgen receptors.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:7211-7215(1988).
RN [3]
RN SEQUENCE FROM N.A.; AND VARIANT TFM GLN-735.
RP MEDLINE-90258822; PubMed-2341409;
RA Yarbrough W.G., Quarby V.E., Simental J.A., Joseph D.R., Sar M.,
RA Lubahn D.B., Olsen K.L., French F.S., Wilson E.M.;
RT "A single base mutation in the androgen receptor gene causes androgen
RT insensitivity in the testicular feminized rat.";
RL J. Biol. Chem. 265:8893-8900(1990).
CC -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: HIGHEST LEVELS IN THE SEMINAL VESICLE,
CC VENTRAL PROSTATE AND COAGULATING GLAND WITH LOWER LEVELS IN THE
CC KIDNEY, AND LEVATOR ANI MUSCLE.
CC -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC -!- DISEASE: Defects in AR are a cause of androgen insensitivity. Rats
CC with this syndrome are called testicular feminized (TFM).
CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC NR3 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; M20133; AAA40733.1; -
CC EMBL; M23284; AAA40759.1; -
CC EMBL; J05454; AAA40734.1; -
CC PIR; A34943; A34943.
CC HSSP; P06536; IGDC.
CC TRANSFAC; T00042; -
CC InterPro; IPR001103; Andrgn_receptor.
CC InterPro; IPR000536; Hormone_rec_lig.
CC InterPro; IPR001628; Znf_C4steroid.
CC Pfam; PF00104; hormone_rec; 1.
CC Pfam; PF00105; zf-C4; 1.
CC Pfam; PF02166; Androgen_recep; 1.
CC PRINTS; PR00047; STROIDFINGER.
CC ProDom; PD000035; Znf_C4steroid; 1.
CC SMART; SM00430; HOLI; 1.
CC SMART; SM00399; Znf_C4; 1.
CC PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger; Steroid-binding; Disease mutation.
FT DOMAIN 1 540
FT DNA_BIND 542 607 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 542 562 C4-TYPE.
FT ZN_FING 578 602 C4-TYPE.
FT DOMAIN 673 902 LIGAND-BINDING.
FT DOMAIN 63 67 POLY-ARG.
FT DOMAIN 174 195 POLY-GLN.
FT DOMAIN 370 376 POLY-PRO.
FT DOMAIN 394 400 POLY-ALA.
FT DOMAIN 444 450 POLY-GLY.
FT VARIANT 735 735
FT
FT
FT CONFLICT 195 195 MISSING (IN REF. 3).
FT CONFLICT 389 389 S -> L (IN REF. 2).
SQ SEQUENCE 902 AA; 43F4064759FDCBD CRC64;
Query Match 85.5%; Score 4200.5; DB 1; Length 902;
Best Local Similarity 84.6%; Pred. No. 3.6e-193;
Matches 796; Conservative 36; Mismatches 53; Indels 57; Gaps 4;
QY 1 MEVQLGLGRVYPRPPSKTYRGAFQNLFSQVREVIONPGPRHPEAASAPPGLSLLLQQQ 60
Db 1 MEVQLGLGRVYPRPPSKTYRGAFQNLFSQVREVIONPGPRHPEAASAPPGLSLLLQQQ 54
QY 61 QQQQQQQQQQQQQQQQQQQQQQSPROQQQQQ-GEDGSPQAHRRGPTGYLVLDDEEQPS 119
Db 55 -----OQRETSPRRRRROQHEDGSPQAHIRGTGYLALEREQPS 96
QY 120 QPQSALECHPERGCVPEPGAANAAKGLPQLPAPDEDDSAAPSTLSLLGPTFFGLSSC 179
Db 97 QQQQASGEHPESGCLPEPGAATAPGKLPQPPAPDDQDSDAAPSTLSLLGPTFFGLSSC 156
QY 180 SADKLDLSEATMQLL-----OQQQEQAYSESGSSSRAREAGAPT 222
Db 157 SADKLDLSEATMQLLQQQQQQQQQQQQQQQQQQQEVISEGSSSVRAREATGAPSS 216
QY 223 SKDNYLGGTSTISDNAKELKAVSVSMGLGVLEALHLSPEQLRGDCMYAPLLGVPPAVR 282
Db 217 SKDNYLGGTSTISDNAKELKAVSVSMGLGVLEALHLSPEQLRGDCMYAPLLGVPPAVR 276
QY 283 PTPCAPLAECGSLDSDSAGKSTEDTAETAYSPFKGGYTKLEGESLGCSSGAAGSSGTLE 342
Db 277 PTPCAPLAECGSLDDEGPGKGTETAYSPFKGGYTKLEGESLGCSSGAAGSSGTLE 336
QY 343 LPSTLSLYSGALDEAAAAQSDRYNFPFLAAGPPPPPPPPHPPHARIKLENLDYGSAAW 402
Db 337 IPSSLSLYSGAVDEAAAAQYQNRDYNFPFLALSGPPPPPPPPHPPHARIKLENLDYGSAAW 396


```
QY 403 AAAACRYGDLASLHGAAGAGSGSPSAASASSWHTLTFTAEGLYGPCGGGGGGG 462
Db 397 AAAACRYGDLASLHGGSVAGSTGSPATASSWHTLTFTAEGLYGP----- 445
QY 463 GGGGGGGGGGGGGGAGAVAPYGYTRPPQGLAGQSDFTAPDVWYPGMVSRVYPSPTC 522
Db 446 -----GGGGGSSSPDAGVAPYGYTRPPQGLASQEGDFSASEVWYPGVWNRVYPSPC 501
QY 523 VKSEMPWMDSYSGPYGDMRLTARDHVLPIIDYEPPOKTCCLICDEASGCHYALTCGS 582
Db 502 VKSEMPWMDSYSGPYGDMRLDSTRDHVLPIDYEPPOKTCCLICDEASGCHYALTCGS 561
QY 583 KVFFKRAAEGKQKYLCAASNDCTIDKFRKNCPSRLKCYEAGMTLGARKLKLGNLK 642
Db 562 KVFFKRAAEGKQKYLCAASNDCTIDKFRKNCPSRLKCYEAGMTLGARKLKLGNLK 621
QY 643 LOEEGFASSTPTTETQKLTSHIEGYEQPIFLNVLAEIPEGVVACAGHNDNQPSFA 702
Db 622 LOEEGFASSTPTTETQKLTSHIEGYEQPIFLNVLAEIPEGVVACAGHNDNQPSFA 681
QY 703 ALLSSINELGERQLVHVVKWAKALPGFRNLHVDDQMAVIOYSWGLMVFAMGWSFTNVN 762
Db 682 ALLSSINELGERQLVHVVKWAKALPGFRNLHVDDQMAVIOYSWGLMVFAMGWSFTNVN 741
QY 763 SMLYFAPDLVFNEMHKSRYSCVVRHLSQEF3WLQITPOEFLCMKALLFSIIPV 822
Db 742 SMLYFAPDLVFNEMHKSRYSCVVRHLSQEF3WLQITPOEFLCMKALLFSIIPV 801
QY 823 DGLKNQKFFDELMMYIKELDRITACKRNPTSCSRRYQTLKLLDSVQPTARELHQFTF 882
Db 802 DGLKNQKFFDELMMYIKELDRITACKRNPTSCSRRYQTLKLLDSVQPTARELHQFTF 861
QY 883 DLLIKSHMVSVDFPPEMAEIIISVQPKILSGKVPIYFHTQ 923
Db 862 DLLIKSHMVSVDFPPEMAEIIISVQPKILSGKVPIYFHTQ 902

RESULT 8
ANDR_MOUSE
ID ANDR_MOUSE STANDARD; PRT; 899 AA.
AC P19091;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Androgen receptor (dihydrotestosterone receptor).
GN AR OR NR3C4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=90386642; PubMed=2403358;
RA He W.W., Fischer L.M., Sun S., Bilhartz D.L., Zhu X., Young C.Y.F.,
RA Kelley D.B., Tindall D.J.;
RT "Molecular cloning of androgen receptors from divergent species with
RT a polymerase chain reaction technique: complete cDNA sequence of the
RT mouse androgen receptor and isolation of androgen receptor cDNA
RT probes from dog, guinea pig and clawed frog.";
RL Biochem. Biophys. Res. Commun. 171:697-704(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91133433; PubMed=2178222;
RA Gaspar M.L., Meo T., Tosi M.;
RT "Structure and size distribution of the androgen receptor mRNA in
RT wild-type and tm/y mutant mice.";
RL Mol. Endocrinol. 4:1600-1610(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=91354214; PubMed=1883336;
RA Faber P.W., King A., van Rooij H.C.J., Brinkmann A.O., de Both N.J.,
```

```
RA Trapani J.;
RT "The mouse androgen receptor. Functional analysis of the protein and
RT characterization of the gene.";
RL Biochem. J. 278:269-278(1991).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=92017874; PubMed=1681426;
RA Charest N.J., Zhou Z., Lubahn D.B., Olsen K.L., Wilson E.M.,
RA French F.S.;
RT "A frameshift mutation destabilizes androgen receptor messenger RNA
RT in the tm mouse.";
RL Mol. Endocrinol. 5:573-581(1991).
CC THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC -1- MISCELLANEOUS: IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS
CC ARE THOUGHT TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS;
CC HORMONE BINDING GREATLY INCREASES RECEPTOR AFFINITY; THE
CC HORMONE-RECEPTOR COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA
CC SEQUENCES UPSTREAM OF TRANSCRIPTIONAL START SITES.
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC NR3 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S56585; AAB19916.1; -
CC EMBL; X53779; CAA37795.1; -
CC EMBL; M37890; AAA37234.1; -
CC EMBL; X59592; CAA42160.1; -
CC PIR; A35895; A35895.
CC PIR; A37255; A37255.
CC PIR; A37908; A37908.
CC PIR; S17198; S17198.
CC PIR; S34398; S34398.
CC HSSP; P06536; LGDC.
CC TRANSFAC; T00041; -
CC MGD; MGI:88064; Ar.
CC InterPro; IPR001103; Andrgn_receptor.
CC InterPro; IPR000536; Hormone_rec_lig.
CC InterPro; IPR001628; Znf_C4steroid.
CC Pfam; PF00104; hormone_rec; 1.
CC Pfam; PF00105; zf-C4; 1.
CC Pfam; PF02166; Androgen_recep; 1.
CC PRINTS; PR00047; STROIDFINGER.
CC ProDom; PD000035; Znf_C4steroid; 1.
CC SMART; SM00430; HOLI; 1.
CC SMART; SM00399; Znf_C4; 1.
CC PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
CC Receptor; Transcription regulation; DNA-binding; Nuclear protein;
CC Zinc-finger; Steroid-binding.
FT DOMAIN 1 537 MODULATING (BY SIMILARITY).
FT DNA_BIND 539 604 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 539 559 C4-TYPE.
FT ZN_FING 575 599 C4-TYPE.
FT DOMAIN 670 899 LIGAND-BINDING.
FT DOMAIN 63 67 POLY-ARG.
FT DOMAIN 174 193 POLY-GLN.
FT DOMAIN 367 373 POLY-PRO.
FT DOMAIN 391 397 POLY-ALA.
FT DOMAIN 441 447 POLY-GLY.
SQ SEQUENCE 899 AA; 98193 MW; FD9EE07C07F7A568 CRC64;
Query Match 85.0%; Score 4176; DB 1; Length 899;
Best Local Similarity 84.1%; Pred. No. 5.2e-192;
```

ANDR_RABIT		STANDARD;	PRT;	709 AA.
ID	ANDR_RABIT			
AC	P49699;			
DT	01-FEB-1996	(Rel. 33, Created)		
DT	01-FEB-1996	(Rel. 33, Last sequence update)		

QY 378 PPPPPHPPHARIKLENPLDYSAWAAAAAQQRYGDLASLHGAGAGPGSGSPSAASSW 437
 Db 181 -QPPPPHPR-RIKLENPLDYSAWAAAAAQQRYGDLASLHGAGAGPGSGSPSAASSW 238
 QY 438 HTLFTAEGQLYPCGG 497
 Db 239 HTLFTTEGQLYGLC-----GGGGGGGPGGAGAVAPYGYTRPPQGLTQGE 283
 QY 498 SDFTAPDVMYVPGMYSRVPYSPCTVKSEMGPMWDSYSGPYGDMRLETARDHVLPIDIYF 557
 Db 284 GDFPAPEWYVPGVSVRYPYFNPSCYKSEMGPMWDSYSGPYGDMRLETARDHVLPIDIYF 343
 QY 558 PPOKTLICGDBASCHYGALTGCSCKVFFKRAAEGKQKYLCAASRNDCTIDKFRKNCPS 617
 Db 344 PPOKTLICGDBASCHYGALTGCSCKVFFKRAAEGKQKYLCAASRNDCTIDKFRKNCPS 403
 QY 618 CRLKRCYEGAGTGLGARKLKLGNLKLQEGEGASSTTPTTETQKLTVSHIEGYECOPIF 677
 Db 404 CRLKRCYEGAGTGLGARKLKLGNLKLQEGEGASSTTPTTETQKLTVSHIEGYECOPIF 463
 QY 678 LNVLEAEPGVVYVAGHNDNQDPSFAALLSSNLGELGQVHVVKWAKALPGERNLHVDDQ 737
 Db 464 LNVLEAEPGVVYVAGHNDNQDPSFAALLSSNLGELGQVHVVKWAKALPGERNLHVDDQ 523
 QY 738 MAVIYQSWMGLVYFANGWRSFTNNSRMVYFAPDLVFNEMRHSKMYOCVVRMHLQSE 797
 Db 524 MAVIYQSWMGLVYFANGWRSFTNNSRMVYFAPDLVFNEMRHSKMYOCVVRMHLQSE 583
 QY 798 FGWLIQITPOEFLCMKALLFSIIPVDGLKNOKFDELRLMNYIKELDRITACKRNPSTCS 857
 Db 584 FGWLIQITPOEFLCMKALLFSIIPVDGLKNOKFDELRLMNYIKELDRITACKRNPSTCS 643
 QY 858 RRFYQLKLLDSVQPIARELHOFTEDLLIKSHMVSVDPEMMAEIIISVQVVKILSKVXP 917
 Db 644 RRFYQLKLLDSVQPIARELHOFTEDLLIKSHMVSVDPEMMAEIIISVQVVKILSKVXP 703
 QY 918 IYFHTQ 923
 Db 704 IYFHTQ 709

RESULT 10
 PRGR RABIT
 ID PRGR RABIT STANDARD; PRT; 930 AA.
 AC P06166;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Progesterone receptor (PR).
 GN PR OR NR3C3.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87067449; PubMed=3538016;
 RA Loosfelt H., Atger M., Misrahi M., Guiochon-Mantel A., Meriel C.,
 RA Logeat F., Benarous R., Milgrom E.;
 RT "Cloning and sequence analysis of rabbit progesterone-receptor
 complementary DNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:9045-9049(1986).
 CC -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
 CC THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
 CC PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 CC NR3 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M14547; AAA314443.1;
 DR PIR; A25923; A25923.
 DR HSSP; P06401; 1A28.
 DR TRANSFAC; T00697;
 DR InterPro; IPR000536; Hormone_rec_lig.
 DR InterPro; IPR000128; Progesterone_receptor.
 DR InterPro; IPR001628; Znf_C4steroid.
 DR Pfam; PF00104; hormone_rec; 1.
 DR Pfam; PF00105; zf-C4; 1.
 DR PRINTS; PR00047; STROIDFINGER.
 DR PRODOM; PD000035; Znf_C4steroid; 1.
 DR SMART; SM00430; ZNF_C4; 1.
 DR SMART; SM00399; ZNF_C4; 1.
 DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
 DR Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KW zinc-finger; Steroid-binding.
 FT DOMAIN 1 565 MODULATING, PRO-RICH.
 FT DNA_BIND 568 633 NUCLEAR RECEPTOR-TYPE.
 FT ZN_FING 568 588 C4-TYPE.
 FT ZN_FING 604 628 C4-TYPE.
 FT DOMAIN 678 930 STEROID-BINDING.
 FT DOMAIN 184 188 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 SQ SEQUENCE 930 AA; 98666 MW; 644FFAC13BF2F883 CRC64;
 Query Match 25.7%; Score 1262.5; DB 1; Length 930;
 Best Local Similarity 34.4%; Pred. No. 2.3e-53;
 Matches 349; Conservative 127; Mismatches 313; Indels 225; Gaps 38;
 QY 39 PRHPEAAGAAPP-----GASLLLLQQQQQQQQQQQQQQQQQQQQQQQQQETS----- 85
 Db 9 PRAPHVAGAPSPTEVGSQLL-----GRDPGPGFGSOTSEASVSAIPIS 55
 QY 86 -----PROQQQQQGEDGSPQ-----AHRGP---TGVL--V 111
 Db 56 LDGLLFRPCQCNPPDGKTDQPPSLSDVEGAFPGVEAGAGDSSRPPEKDSGLDSV 115
 QY 112 LDEEQPSQP---QSALECH-----PERGCVPEGAAGAASGLPQPLPAPDE- 157
 Db 116 LDTLLAPGSGGSHASPATCEALSPWCLFGPD---LPEDPRAAPATKGLVAPLMSRPDK 172
 QY 158 --DDSA-----APSTLSLL---GPTFPGL-----SSCSADLKDL-SEAT 192
 Db 173 AGDSSGTAAHKVLPRGLSPSRQLLLPSSGSPHPVAPKVPQPAVQVDEDSSESGTV 232
 QY 193 MQLLQQQQQEAIVSESSSGRAREASGAPTSDKNYLGSTTI--SDNAKELCKAVSVSMGL 251
 Db 233 GPLLKGPFRALGTAAGGAAPVAGAA-----GGVALVPKEDSRFSAPRVSLA--- 282
 QY 252 GVEALEHLSGCEQLRG-----DCMYAPILGVPPAVRPPCAPLAECKSLDDSGKST-- 305
 Db 283 --EQDAPVAPGRSPLATSVVDVFIHVPIPLNHPATFATRTQLLE--GESYDGGAAASPF 338
 QY 306 -----ETAEYSPFKGG-----YTKLEGESLGCSSAAAGSGTLELPSTLSLYKSGA 354
 Db 339 VPQRGSPSASSTPVAGGDFPDCTYPPDAEPK-----DDAFPLYGDFQ--PPALKIREEE 391
 QY 355 LDEAAAYQSRDY-----NFP--LALAGPPPPPPPPHARIKLENPLDYGSAWAAAA 406
 Db 392 AAEAAASRPTLYVAGANPAAFDFQLAAPPPLPSP-----RVPSRP---GEA----- 438
 QY 407 QCRYGLDIALHGAGAGPGSGSPSAASSSWH---TLFTAE-----EGLYG--PCGGGG 457
 Db 439 -----AVAASPGSASVSSSSSGSTLECLIKYKAGAPPOQGFAPLPCPPGA 486
 QY 458 GGGGGGGGGGGGGGGGGGGGAGAVAPYGYTRPPQGLAGQESDFTAPDVMYVPGMWSR--- 514

DB 487 GACLLPRDLPTSTSAAGAAPALY-----PTLGLNG-----LPOLCQAVALKEGLP 536

QY 515 ---VPYSPPTCVKSEMPWMDSYSGPYGDMRLTARDHVLPIDYYP--PQKTCICGDE 569

DB 537 QVITYP-----LNYLRPDSEASQSP-----QYSFESLPQKICICGDE 574

QY 570 ASCHYGALTCGCKVFFKRAAEKQKYLCAASNDCTIDKFRKNCPCSLRKLKYEAGWT 629

DB 575 ASCHYGVLTCGCKVFFKRAEGBQNYLCAGNDCIVDKIRKNCPCACLRKRCQAGMV 634

QY 630 LGARKLKLNLKLOEGEASSTPT--EETQKLTVSHIEGYEQPIFLNVLAEIPG 687

DB 635 LGRKFKFKFKVVRALDAVALPQVGPINESQRTFSPSQEIQLPLINLLMSIEDP 694

QY 688 VVACAGHNNQDPSFALLSLNGLERQLVHVVKAKALPGFRNLHVDQMAVQYSWNG 747

DB 695 VIVAGHNTKPTDTSSTSLTSLNQLGERQLLSVVKVSKSLPGFRNLHVDQITLIQYSWS 754

QY 748 LMFVAMGRSFTNVNSMLYFAPDLVNEVMEKSRWSQVVRMHLISOEFGWLQITPOE 807

DB 755 LMFVGLGRYSKYHVGOMLYFADPLINEQRMKESFYSLCTMWOIPOEFVKLYQSDE 814

QY 808 FLCKMALLFSIIPVGLKQKFFDELRLMYIKELDRITIAKRNKPTCSRFRYQTLKLL 867

DB 815 FLCKMALLLTITPLEGRSQSQPEEMRSSYIRELIKAIGLRQKGVVSSQSRFYQTLKLL 874

QY 868 DSVQPIARELHOFTDLIKSHMVSVDFPEMAEIIISVQVVKILSGKVKPIYFH 921

DB 875 DNLHDLVKQLHLYCLNTFIQSLRSLSVFPEMMSEVIAAQLPKILAGMKVLPFLH 928

RESULT 11

ID PRGR_HUMAN STANDARD; PRT; 933 AA.

AC P06401; Q9UPF7;

AT 01-JAN-1998 (Rel. 06, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Progesterone receptor (PR).

GN PGR OR NR3C3.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.

RX MEDLINE-90228361; PubMed-2328727;

RA Kastner P., Krust A., Turcotte B., Stroop U., Tori L., Gronemeyer H., Chambon P.;

RT encoding the two functionally different human progesterone receptor forms A and B.;

RT EMBO J. 9:1603-1614(1990).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE-87184565; PubMed-3551956;

RA Misrahi M., Atger M., D'Auriol L., Loosfelt H., Meriel C., Fridlansky F., Guiochon-Mantel A., Galibert F., Milgrom E.;

RT "Complete amino acid sequence of the human progesterone receptor deduced from cloned cDNA.;"

RT Biochem. Biophys. Res. Commun. 143:740-748(1987).

RN [3]

RP SEQUENCE FROM N.A.

RA Kieback D.G., Aguilnik I.U., Tong X.-W.;

RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.

RN [4]

RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 682-933.

RX MEDLINE-98282128; PubMed-9620806;

RA Williams S.P., Sigler P.B.;

RT "Atomic structure of progesterone complexed with its receptor.;"

RL Nature 393:392-396(1998).

CC -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN

CC THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.

CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A AND B (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN, A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.

CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.

CC NR3 SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by, and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

CC EMBL; X51730; CAA36018.1; -

DR EMBL; M15716; AAA60081.1; -

DR EMBL; AF016381; AAD01587.1; -

DR PIR; A03245; QRHUP.

DR PIR; S09971; S09971.

DR PDB; 1A28; 15-JUL-98.

DR TRANSFAC; T00696; -

DR Genew; HGNC:8910; PGR.

DR MIM; 264080; -

DR InterPro; IPR000536; Hormone_rec_lig.

DR InterPro; IPR000128; Progesterone_receptor.

DR InterPro; IPR001723; Steroid_receptor.

DR InterPro; IPR001628; Znf_C4steroid.

DR Pfam; PF00104; hormone_rec; 1.

DR Pfam; PF00105; zf-C4; 1.

DR Pfam; PF02161; Prog_receptor; 1.

DR PRINTS; PR00398; STRDHORMONER.

DR PRINTS; PR00047; STROIDFINGER.

DR ProDom; PD000035; Znf_C4steroid; 1.

DR SMART; SM00430; HOL1; 1.

DR SMART; SM00399; Znf_C4; 1.

DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.

DR Receptor; Transcription regulation; DNA-binding; Nuclear protein; Zinc-finger; Steroid-binding; Phosphorylation; Alternative splicing; 3D-structure.

KW DOMAIN 1 566 MODULATING, PRO-RICH.

FT DNABIND 567 632 NUCLEAR RECEPTOR-TYPE.

FT ZNFING 567 587 C4-TYPE.

FT ZNFING 603 627 C4-TYPE.

FT DOMAIN 681 933 STEROID-BINDING.

FT DOMAIN 183 187 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

FT MOD_RES 41 41 PHOSPHORYLATION (BY CK).

FT MOD_RES 227 227 PHOSPHORYLATION (BY CK).

FT MOD_RES 232 232 PHOSPHORYLATION (BY CK).

FT MOD_RES 552 552 PHOSPHORYLATION (BY CK).

FT MOD_RES 793 793 PHOSPHORYLATION (BY CK).

FT VARSPLIC 1 164 MISSING (IN ISOFORM A).

FT CONFLICT 226 226 G -> S (IN REF. 1).

FT CONFLICT 256 256 V -> S (IN REF. 1).

FT CONFLICT 344 344 S -> T (IN REF. 1).

FT CONFLICT 660 660 L -> V (IN REF. 2).

SQ SEQUENCE 933 AA; 98994 MW; 80414B/F1F317F8E CRC64;

Query Match 25.6%; Score 1256; DB 1; Length 933;

Best Local Similarity 34.5%; Pred. No. 4.7e-53;

Matches 349; Conservative 119; Mismatches 329; Indels 214; Gaps 33;

QY 38 GPRHPEAASAP-----PGA-----SLLLLLQOOOQOOOQ 66

DB 8 GPRHPEAASAP-----PGA-----SLLLLLQOOOQOOOQ 67

QY 67 QOOOQOOOQOOOQOOOQOOOQOOOQOOOQOOOQOOOQOOOQ 124

DB 68 DPSEKTDQDOOQSLDVEGAYSRAEATRGAGSSSSPPEKDSGLDVLTLAPSGPGS 127

QY 125 LECHPERGCVPEFGAASVSKG--LPQOLPAPPPDEDDSAAPSTLSLIGTFPP-----GL 176
 Db 128 -----QPSPPACEVTSWCLRGPELP-----EDPAATQATQVLSPLMSRSGCKVGD 174
 QY 177 SSCSADLKILSE--ASTWQLL-----QQQOQAEVSGSSSGRAREASGA 219
 Db 175 SSGTAAAHKVLPRGLSPARQQLLLPASESPHWSGAPVKPSPQAAAIVEEEDGSESESAG 234
 QY 220 P-TSSKDNVYLGTSYISDNA-----KELCKAVSYSMGLGVLEAHLSPGE 263
 Db 235 PLLKCKPRALGGAAGGAAVPPGCAAGGVALVPKEDSRFSAPRVAL-VQDDAPMAPGR 293
 QY 264 QLRG-----DCMYAPLLGVPPVAPRTPCAPLAECCKSLDSDSAGKSTEDTAESYSPFKGYT 319
 Db 294 SPLATTVMDFIHVPIL-----PLNHALLAARTQLEDES-----YDGGAGAA 337
 QY 320 KGLESGICSGS--AAAGSSGTLELP-----STLSLY-----KSGALDE 357
 Db 338 AFAPRRSPCASSTPVAVGDFPDCAYPDABPKDDAYPLYSDFPPALKIKEEEGA--E 395
 QY 358 AAAYOSRDY-----NPLALAGPPPPPPPPHPPHARIKLENPLDYGSAAWAAAAQ 407
 Db 396 ASARSPSYLVAGANPAAPDFPL--GPPPLPP-----RATP 431
 QY 408 CRYGDLASLHGAAGPGSGSPSAASSWH---TLFTAE-----EQQLYGPCGGGGGG 459
 Db 432 SRPGE-----AAVTAAPASVSSASSSGSTLECLYKAEGAPPOGQFPAPPCAPCAS 486
 QY 460 GGGGGGGGGGGGGGAGAVAPVGYTRPPQGLAGQESDFTAPDVWYPGMWSR-VPYP 518
 Db 487 GCLLPDGLPSTSASAAAGA-AFALY--PALGLNG-----LPQLGYQAALVKEGLUPQV 537
 QY 519 SPTCVKSEMPWMDSYSGPYGDMRLTARDHVLPIDYFP--POKTLICGDEASGCHYG 576
 Db 538 YPPYL-NYLRPDSASQSP-----QYSESLPQKICLICGDEASGCHYG 580
 QY 577 ALTGSCKVFVKRAEKGKQKYLCAASRNDICDKFRKNCPSCLRKYEAQWGLGARKLK 636
 Db 581 VLTGSKVFVKRAEKGHNYLCAGRNDICVDKIRKNCACLRKCCQAGWVLGGRKFK 640
 QY 637 KGLNKLQEGEASSTSP-----TEETQKLTVSHIEGYEQPIFLNLEAIEPGVVC 690
 Db 641 KFNKVRVRLDALVALPQLGVNPNESQALSORFTFSPQDIQLIPPLINLMSIEPDVY 700
 QY 691 AGHDNVPDFAALLSSELGERQLVHVVKAKALPGFRNLHVDQMAVIOYSWMLGMV 750
 Db 701 AGHDNTKPDTSSTSLTSLNOLGERQLSVVVKWSKSLPGFRNLHVDQITLIQYSWMLMV 760
 QY 751 FAMGWRSTFVNSRMLYFAPDLVFNVEYRHKSRMYSCVVRHLSQEFGLQITPQEFLLC 810
 Db 761 FGLGWSYKHKVSGMLYFAPDLILNEORMKESFYSLCLTMWQIPQEFVQLQVSEFFLC 820
 QY 811 MKALLFSIIPVGLKNQKFFDELRMNYIKELDRITACKRKNPTSCSRRYQLTKLDSV 870
 Db 821 MKVLLNTIPLGLRSQTQFEMRSSYIRELIKAICLRQKGVVSSORFYQITKLIDLNL 880
 QY 871 OPTARELHQFTDLLIKSHMVSDVDFPMMAEIIISVQPKILSKVKPIYFH 921
 Db 881 HDLVKQLHLXCLNTFIQSRALSVFEPPEMSEVIAAQLPKILAGWVKPPLLFH 931

RESULT 12
 PRGR_RAT
 ID PRGR_RAT STANDARD; PRT; 923 AA.
 AC Q63449;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Progesterone receptor (PR).
 GN PR OR NR3C3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

NCBI_TaxID=101116;
 [1]
 SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Placenta;
 RX MEDLINE=94130817; PubMed=8299566;
 RA Park-Sarge O.K., Mayo K.E.;
 RT Regulation of the progesterone receptor gene by gonadotropins and
 RL cyclic adenosine 3',5'-monophosphate in rat granulosa cells,";
 Endocrinology 134:709-718(1994).
 CC -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
 CC THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
 CC PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 CC NR3 SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L16922; AAA19916.1;
 DR HSP; P06401; I428.
 DR TRANSFAC; T04681;
 DR InterPro; IPR000536; Hormone_rec_lig.
 DR InterPro; IPR00128; Progesterone_receptor.
 DR InterPro; IPR001628; ZnF_C4steroid.
 DR Pfam; PF00104; hormone_rec; 1.
 DR Pfam; PF00105; zf-C4; 1.
 DR Pfam; PF02161; Prog_receptor; 1.
 DR PRINTS; PR00047; STROIDFINGER.
 DR ProDom; PD000035; ZnF_C4steroid; 1.
 DR SMART; SM00430; HOLI; 1.
 DR SMART; SM00399; ZnF_C4; 1.
 DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
 KW Receptor; transcription regulation; DNA-binding; Nuclear protein;
 KW Zinc-finger; Steroid-binding.
 FT DOMAIN 1 556 MODULATING, PRO-RICH.
 FT DNA_BIND 557 622 NUCLEAR RECEPTOR-TYPE.
 FT ZN_FING 557 577 C4-TYPE.
 FT ZN_FING 593 617 C4-TYPE.
 FT DOMAIN 671 923 STEROID-BINDING.
 FT DOMAIN 184 188 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 SQ SEQUENCE 923 AA; 99407 MW; 05384B9656BF22DC CRC64;
 Query Match 25.1%; Score 1234; DB 1; Length 923;
 Best Local Similarity 33.1%; Pred. No. 5.2e-52;
 Matches 344; Conservative 129; Mismatches 299; Indels 266; Gaps 39;
 QY 15 PSKTYRGA-----FQNLFSQVREVIQNGPRHPEAASAPP---GASLLLLQOQQ 61
 Db 19 PSPTHGSPLLARLDPDPEQ-----GQSHSDASSVVSPIISLDRLLFSSRC 65
 QY 62 QQQQQQQQQQQQQQ-----QQQQQSTSPROQQQQQGGDSGFOAHRGPTYL 110
 Db 66 QAQELPDEKTNQQSLSVDEGAFSGVEASRRSRNRPAPRPEKDSRLDLS----- 113
 QY 111 VLDEQQPQSQPSQSALECHPERGC-----VPEPGAAVAASGLPQOLPAPDPE- 157
 Db 114 VLDTLAPSGPEQSTSP--ACEAITSWCLFGLPELDPEDPSVPATKGLSPLMSRPEK 171
 QY 158 ---DSSAA-----PSTLS---LIGPTF-----PGLSSCSADLKD---ILSEA 190
 Db 172 AGDSSTGAGQKVLKAVSPRQLLLPTSGSAHWPGAGVKPSQQAIVEEEDGLETEG 231
 QY 191 STMQLLQOQQQEAIVEGSSSSGRAREASCAPTSSKDNVILGGTSTI--SDNAKELKAVSM 249
 Db 232 SAGPLLKSRALRGMCSSGGGVATANAPGAP-----GGVTLPVKEDSRFSAPRVSLQ 284

307	QY	DTAEYSPFKGGYTKGLESGSGSAAAGSSGTLELPSTLSLYKSGALDEAAAYSRDY	369
197	Db	-----GPGEGL-----APAAAAPRAVE-----PGA-----GQDY	222
367	QY	YNFPLALAGPPPPPPHPPHARIKLENPLDYSAAWAAAAAQCRYGDLASLIAGAGAPGS	426
223	Db	LHVPI-----LPLNSAFLASRTR-QLLDVEAYADGSAFGPRS	258
427	QY	GSPSAAASSWHTLFTAEAGOLYPCPGGGGGGGG-----GGGGGGGGGGGG	475
259	Db	-SPSVPAAD-----LAEYG--YPPDPCKEGPAYEGFQSALKIKEEGVGLPAAPPFFLG	309
476	QY	GEAGAVAPYGYTRPPQGLAGQESDFTA-----PDVWYEGGMVSRVPYPS-----	519
310	Db	AKA--APADFAOPPR--ACGEPSLECVLYKAEPPLLPAGYGPAAAPDSLSTSAAPPGL	364
520	QY	-----PTCVKSEGMPPWMDSTSYGPI--GDMRLLETARDHVLPIDYTFPQKTC	563
365	Db	YSPGLNGHHQALGFPAAVLKLEGPLQ-----CPPYLGYVRPDTTETSSOQYSFESLPQKIC	421
564	QY	LICGDEASGCHYGALTCGSKCVFFKRAAECKQKYLCASRNDCTIDKFRKNCPCSLRKRC	623
422	Db	LICGDEASGCHYGVLTCGSKCVFFKRAEGCHNYLCAGRNDCIVDKIRKNCFACLRKC	481
624	QY	YEAGMTLGARKLKLGNLK-----LQEBGEASSTSPTEETQKLTVSHIE	659
482	Db	CQAGMWLGRRKFKLNMKVVRTLDAVLQOPAVLQDE-----TQSLTQRLSFPNQ	532
670	QY	GYEQPTFLNVLNLEAIEPGVVCAGHDNNQPSFALLSSLNELGERQLVHVVKAKALPGF	729
533	Db	EIPFVPMISVLRGIEPEVYAGVDNTPKPTPSSLLTSLNHLGERQLLCVYKWSKLLPGF	592
730	QY	RNLHVDDQMAVIOYSWGLMVFAMGWRBFTNVNSRMLYFAPDLVFNERYMHKSRMYSOVY	789
593	Db	RNLHIDQITLIIQYSWMLMVFAMGWRBFTNVNSRMLYFAPDLVFNERYMHKSRMYSOVY	849
790	QY	RMRLHSQFQWLQITPQEFELCMKALLFSIIPVDGLKNQKFEFDELRMNYIKELDRIIACK	899
653	Db	SMWQLPQEFVRLQVQSEEFELCMKALLLNTIPLGLRSQSFDEMTSYIRELVKAIGLR	712
850	QY	RKNPTCSRRFYQLTKLSDSQPTARELHOFTEPDLLIKSHWVSDPPEMMAELISVQVPK	909
713	Db	QKGVANSQRFYQTLKMSDMSHDLVKQLHLFCLNTPLQSRALSVFEPPMSEVIAAQLPK	772
910	QY	ILSGKVRPIYFH 921	
773	Db	ILAGWVAPLLEFH 784	
RESULT 14			
PRGR_MOUSE			
ID	PRGR_MOUSE	STANDARD;	PRT; 923 AA.
AC	Q00175;		
DT	01-DEC-1992 (Rel. 24, Created)		
DT	01-DEC-1992 (Rel. 24, Last sequence update)		
DT	15-JUL-1999 (Rel. 38, Last annotation update)		
DE	Progesterone receptor (PR).		
GN	PR OR NE3C3 OR PR		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=91299759; PubMed=2069958;		
RA	Schott D.R., Sinyamala G., Schneider W., Parry G.;		
RT	"Molecular cloning, sequence analyses, and expression of		
RT	complementary DNA encoding murine progesterone receptor.";		
RL	Biochemistry 30:7014-7020(1991).		
RN	[2]		
RP	SEQUENCE OF 1-9 FROM N.A.		
RX	STRAIN=129/SV;		
RC	MEDLINE=95100931; PubMed=7802637;		

RA Hagiwara K., Wu-Peng X.S., Funabashi T., Kato J., Pfaff D.W.:
 RT "Nucleic acid sequence and Dnaase hypersensitive sites of the 5'
 RL Biochem. Biophys. Res. Commun. 205:1093-1101(1994).
 CC -1- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
 CC THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
 CC PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 CC NR3 SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M68915; AAA39971.1; -
 DR EMBL; U12644; AAA66067.1; -
 DR PIR; A39596; A39596.
 DR HSP; P06401; 1A28.
 DR TRANSFAC; T04680; -
 DR MGD; MGI:97567; Pgr.
 DR InterPro; IPRO00536; Hormone_rec_lig.
 DR InterPro; IPRO01128; Progesterone_receptor.
 DR InterPro; IPRO01628; Znf_C4steroid.
 DR Pfam; PF00104; hormone_rec; 1.
 DR Pfam; PF00105; zf-C4; 1.
 DR Pfam; PF02161; Prog_receptor; 1.
 DR PRINTS; PR00047; STROIDFINGER.
 DR ProDom; PD000035; Znf_C4steroid; 1.
 DR SMART; SM00430; HOLI; 1.
 DR SMART; SM00399; Znf_C4; 1.
 DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
 DR Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KW Zinc-finger; Steroid-binding.
 KW DOMAIN 1 556 MODULATING, PRO-RICH.
 FT DNA_BIND 557 622 NUCLEAR RECEPTOR-TYPE.
 FT ZN_FING 577 C4-TYPE.
 FT ZN_FING 593 617 C4-TYPE.
 FT DOMAIN 671 923 STEROID-BINDING.
 FT DOMAIN 184 188 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT SEQUENCE 923 AA; 99073 MW; 9415F1ED34BEE3F CRC64;
 Query Match 24.5%; Score 1205; DB 1; Length 923;
 Best Local Similarity 32.2%; Pred. No. 1.3e-50;
 Matches 337; Conservative 119; Mismatches 301; Indels 288; Gaps 36;
 QY 35 QNCPGPRHPEAASAPP--GASLLLLQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ-----86
 DB 7 KDFQVLTGASGSPPHIGSPLL-----ARLDGPGFGSQHSDVSSVWSPISLD 57
 QY 87 -----RQOQQQGGDGSQA-----HRRGPTY-----LVLD 113
 DB 58 GLLPFRSCRGPELPDGTGQDQSLSDVEGAFSGVEATHREGGRNSPPKDSRLDSDVLD 117
 QY 114 EEOQPSQPSALECHPERGCVPEFGAASKGLPOLPAPPEDDSAAPSTLSLIGPTF 173
 DB 118 SLTPSPGE---QSHAS---PPACEAITSWCLFGPELP-----EDPRSPVATKGLLSP--164
 QY 174 PGLSSCSADKLDLSEASTWMLQQQQQEAIVEGSSGSGRARSAGTSSKKNYLGST 233
 DB 165 -----LMSRPEIKVGDQSGTGRGQ-----183
 QY 234 ISDNAKELKAVSMGLGLVALEHLSPEQL---RGDCMYAPLLGVPVAVRPTCAPLA 290
 DB 184 -----KVLPGK-----LSPRQLLLPTSGSAHPGAGVPS--PQFAA---219
 QY 291 ECKGSLDSDSAGKSTEDTAEPSPFKGYTKGLEGESLG---CSGSAAGSGSCTLELPSTL 347

DB 220 ---GEVEEDS-GLETEGSA---SPLLKSPRALEGTGGGGVAAANAPSAPGGVTLVPKRED 273
 QY 348 SLYKSGALD-----EAAAYQSRDYNNFPL-----A 372
 DB 274 SRFSAPRVSLQDSPAIGRSPPLATTVVDFIHPILPLNHALLAARTFQLLEGESYDGA 333
 QY 373 LAGPPPPPPPP-----HPHARTIKLEN-----PLDYSANAAAAACRYGDLA 414
 DB 334 TAGPPCPSPSPAPSTPVPGRDFDCTVPLEGDKPKEDVFP-----YGDQF 379
 QY 415 S-----LHCAGAGPCGSSP-----SAAASSSWHTLFTAEEGOLYPCGCGGGGGGG 462
 DB 380 TPLGKIKEEGADAA---VRSPPRYLSAGASSSTFPDPLAPAPQAAPSRPGEAAVAG 436
 QY 463 GGGGGGGGGGGGGA-----CAVAPYGYTRPP-----QGLAGQESD 499
 DB 437 GPSSAAVSPASSGSALECLYKAEAPPTQGSFAPLP-CKPPAAASCLLPDPSLPAPAGT 495
 QY 500 FTAPDVWYPGMVKSRVPYSPCTCVKSEMPWMDSYSGPYGDMRLTARDHVLV--IDYFF 557
 DB 496 AAAPAIYQPLGL-NGLP-----QLG-----YQAAVLKDSLPOVYPYPLNYLR 536
 QY 558 P-----POKTCICGDEASGCHYGALTCGCKVFEKRAAEGKQKYLCA 602
 DB 537 PDSEASQSPQYGFDSLPOKICLICGDEASGCHYGVLTCGCKVFEKRAAEGKQKYLCA 596
 QY 603 NDCTIDKFRKNCPSCLRKCYEAGMTLGARKLKLGNLKLQE--EGEA--SSTTSPT 658
 DB 597 NDCIVDKIRKNCPCARLKKCCQAGMVLGGRKFKKVKVRVMTLDGVALPQSVGLPNE 656
 QY 659 --TQKLTVSHIEGYECOPIFLANVLEAIEPGVVCAGHDNQPDFAALLSLNELGERQL 716
 DB 657 QALSQRTFSPNQEIQLVPLINLLMSTEDVIYAGHDNTPKDTSSLLTSLNGLERQL 716
 QY 717 VHYVKAALPGFRNLHVDQMAVIOYSWGLVAFAMGWSRFTVNSRMLYFADPLVNE 776
 DB 717 LSVKWSKSLGFRNLHDDQITLIQYSWMSLWVGLGWSYKHSQOMLYFADPLVNE 776
 QY 777 YRMHKSRYSCVVRMHLQSOEFGWLQITPOEFLCMKALLLSIIPVDGLKNQKFFDELM 836
 DB 777 QRMKELSYSLCTMWWQIPQEFVQLQVTHEEFLCMKVLILLNTIPLEGLRSQSOFEMRS 836
 QY 837 NVIKELDRIIACKRKNPTSCSRREYVOLTKLDSVQPTARELHQFTDLLLKSHMVSDPE 896
 DB 837 STIRELKAIGLRQGVVPTSORFYQLTKLSDLHDLVKQLHLYCLNTFIQSRRTJAVEFP 896
 QY 897 EMMAEIIISVQVPKILSGKVKPIYFH 921
 DB 897 EMMSEVIAAQLPKILAGMVKPLLFH 921
 RESULT 15
 PRGR_SHEEP
 ID PRGR_SHEEP STANDARD; PRT; 377 AA.
 AC Q28590;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE progesterone receptor (PR) (Fragment).
 GN PGR OR NR3C3.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 RN NCBI_TaxID=9940;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Romanov; TISSUE-Uterus;
 RA Medigou T., Tiffoche C., le Gal F., Pelletier J., Thieulant M.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
 CC THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR

```

CC PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC NR3 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z66555; CAA91447.1; -.
DR HSP; P06401; 1A28.
DR InterPro; IPR000536; Hormone_rec_lig.
DR InterPro; IPR00128; Progesterone_receptor.
DR InterPro; IPR01628; Znf_C4steroid.
DR Pfam; PF00104; hormone_rec; 1.
DR Pfam; PF00105; zf-C4; 1.
DR Pfam; PF02161; Prog_receptor; 1.
DR ProDom; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOL1; 1.
DR SMART; SM00399; Znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger; Steroid-binding.
FT NON_TER 1
FT DOMAIN <1 15 MODULATING, PRO-RICH.
FT DNA_BIND 18 83 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 18 38 C4-TYPE.
FT ZN_FING 54 78 C4-TYPE.
FT DOMAIN 128 >377 STEROID-BINDING.
FT NON_TER 377
SQ SEQUENCE 377 AA; 42904 MW; 3141B65587F7493C CRC64;

Query Match 22.7%; Score 1116; DB 1; Length 377;
Best Local Similarity 55.2%; Pred. No. 8,8e-47;
Matches 201; Conservative 73; Mismatches 84; Indels 6; Caps 1;

QY 559 POKTCLICGDEASGCHYGALTCGCKVFFKRAEGKQKYLCAASRNDCTIDKFRKNCPC 618
Db 14 POKICLICGDEASGCHYGALTCGCKVFFKRAEGKQKYLCAASRNDCTIDKFRKNCPC 73
QY 619 RLKCYEAGMTLGARKLKLGNLKLQEGEASSTSP-----TEETQKLTWSHIEGYE 672
Db 74 RLKCCQAGMVLGGKFKFKRVMTLDVALPQVGPINPESQALSQRITFSPSODLQ 133
QY 673 COPIPLNLEAIEPGVVCAGHDNDPDFAALLSSLNGLEROLVHVYKAKALPGFRNL 732
Db 134 LIPPLINLLMSIEDPMVYAGHDNSKPDTSLLTSLNQLGERQLSVYKWSKSLPGFRNL 193
QY 733 HVDDQMAVIOYSWMLGMVAFMGWSFTNVNSRMLYFAPDLVFNRYMHKSRMYSCQVMR 792
Db 194 HIDDQITLIQYSWMLGMVFLGWSYKHSVQMLYFAPDLINBQRMKESFYSLCITMW 253
QY 793 HLSQEFGLQITPQEFCLMKALLFSLIPVDGLANQKFFDELNNYIKELDRITACKRN 852
Db 254 QIPQEFVKLVQSVQEEFLCMKVLLINTIPLEGLRSNQFQEMRSSYIRELIKAIGLRQK 313
QY 853 PTSCSRREYQLTKLLDSVPIARELHOFTFDLLIKSHMVSVDPEMMAEIIISQVPKILS 912
Db 314 VVPSQRYQLTKLLDNLHDLVQLHLYCLNTFTIQSRALSVEFFPEMMSEVIAAQLPKILA 373
QY 913 GKVK 916
Db 374 GMVK 377

```

